PSYC 331 STATISTICS FOR PSYCHOLOGIST

Session 2 – INTRODUCTION TO THE GENERAL STRATEGY OF INFERENTIAL STATITICS

Lecturer: Dr. Paul Narh Doku, Dept of Psychology, UG
Contact Information: pndoku@ug.edu.gh

UNIVERSITY OF GHANA
College of Education
School of Continuing and Distance Education
2014/2015 – 2016/2017

Session Overview

- This session builds upon the first session and provides further insight into some basic statistical concepts that will facilitate the understanding of statistical tests. The goal of this session is to equip students with the ability to:

- **Explain** how the simple toss of a coin can be used to illustrate the concept of probability, and how the probability of independent events is related to the decisions that we take when we are conducting research;
- **Estimate** how the areas under the standard normal curve distribution are related to the decisions taken when we are conducting research. How stated hypotheses are tested with certain probabilities and lastly
- **Describe** how to test a one-tailed and two-tailed hypothesis in research.

Session Outline

The key topics to be covered in the session are as follows:

- A ƌeǀieǁ of iŶfeƌeŶtial statistiĐs aŶd aŶ aŶalogLJ ǁith the ĐoiŶ tossiŶg experiment
- The pƌoďaďilitLJ of iŶdepeŶdeŶt eǀeŶts, the ĐoiŶ tossiŶg edžpeƌiŵeŶt and the 0.05 decision rule
- The staŶdaƌd eƌƌoƌ of the ŵeaŶ aŶd the staŶdaƌd Ŷoƌŵal Đuƌǀe distribution
- HLJpothesis testiŶg
- Tǁo- and one-tailed tests of significance

UNIVERSITY OF GHANA

Reading List

- Opoku, J. Y. (2007). Tutorials in Inferential Social Statistics. (2nd Ed.). Accra: Ghana Universities Press. *Pages 23 – 44*
- MiĐhael Baƌƌoǁ, StatistiĐs foƌ EĐoŶoŵiĐs, AĐĐouŶtiŶg aŶd BusiŶess Studies, ϰth Edition, Pearson
- R.D. MasoŶ , D.A. LiŶd, aŶd W.G. MaƌĐhal, StatistiĐal TeĐhŶiƋues iŶ BusiŶess aŶd EĐoŶoŵiĐs, бϬth Edition, McGraw-Hill

TOPIC ONE: A REVIEW OF INFERENTIAL STATISTICS AND AN ANALOGY WITH THE COIN TOSSING EXPERIMENT

- Inferential statistics: concerned about techniques used in drawing inferences or conclusions from a sample to a population.
- **A** primary goal of inferential statistics is to find out the probability or chance that certain kinds of results based on a sample come from certain population conditions.
- For example, suppose we know that the mean age of a total student population of 20,000 is 20 years (N = 20,000 ; = 20 years) and we take a random sample of size 500 from this population and find that the sample mean is 22 years (n = 500,

= 22 years). There is a difference of 2 years between the sample mean and the population mean.

- We ask the ƋuestioŶ: **What is the probability, or what are the chances that a sample with a mean of 22 years, is indeed coming from a population whose ŵeaŶ is 20 years?”**

A Brief Review Of Inferential Statistics and An Analogy With The Coin Tossing Experiment: cont

- If the sample was randomly drawn from the population and therefore coming from the population, then the difference of 2 years observed between the sample **mean** and population **mean** could have simply been due to chance or error in sampling.

- Inferential statistics will enable us to **calculate the probability** that this difference of 2 years between the sample and population mean is due to chance or not.

- Later in this session you will understand that if the **probability calculated is less than or equal to a certain value**, then any difference that we observe between a sample mean and a population mean could not have arisen from chance.

An illustration with the coin tossing experiment

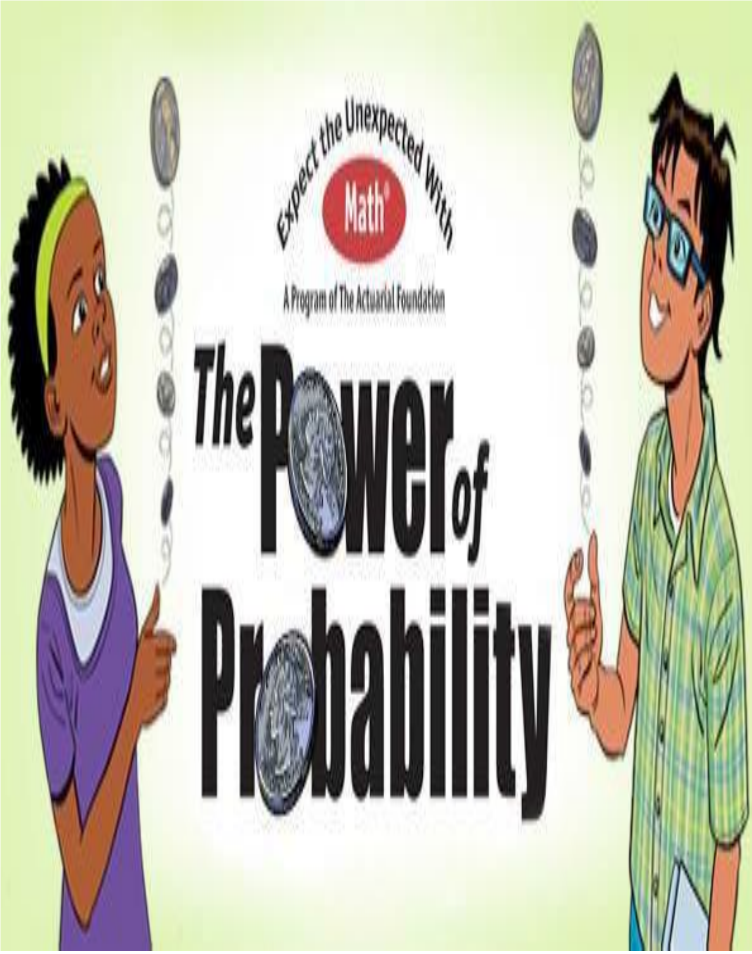
- A coin tossing experiment may be used to explain the general strategy that we use in inferential statistics.
- Consider a coin tossing game with a trusted friend. Your friend is tossing the coin and you both agree before the start of the game that anytime the coin falls heads up, LJou ǁiŶ GH¢ϭ.ϬϬ ;oŶe GhaŶa Đedi aŶd aŶLJtiŵe it falls tails up, LJouƌ friend wins GH¢1.00.
- Assume that after four tosses of the ĐoiŶ, it has falleŶ tails up oŶ eaĐh toss, which means that you have lost GH¢4.00.

An illustration with the coin tossing experiment

You may wonder whether the tossing of the coin is fair or not. Thus either your friend is cheating or you are just being unlucky- four tosses of the coin represents only a small sample. You decide on the latter and continue playing.

After another four tosses (bringing the total number of tosses to eight), your friend still continues to win and you have lost GH¢8.00.

You theŶ deĐide: This is eŶough.   I aŵ Ŷot going to continue playing any longer, for either the tossing is biased and I am being cheated by ŵLJ fƌieŶd, oƌ todaLJ is Ŷot ŵLJ luĐkLJ daLJ.



Decision taking in the coin tossing experiment

- Taking any of the above two decisions means that in effect, you have put forward two hypotheses (hypotheses, because you are not sure which of the two situations above is correct).

- FiƌstlLJ, LJou haǀe put foƌǁaƌd the hLJpothesis that: The ĐoiŶ is uŶfaiƌ; I aŵ just ďeiŶg uŶluĐkLJ eǀeŶ afteƌ eight tosses.

- SeĐoŶdlLJ, LJou haǀe put foƌǁaƌd the hLJpothesis that: The tossing of the coin is indeed biased and therefore I am

Decision taking in the coin tossing experiment

- ***Although you can take a decision to stop playing, you can never be sure that the decision you have taken is absolutely correct.***
- ***Whether you decide to continue or stop playing, there is always the likelihood that your decision could be wrong.***

- The only way that you can be sure that you have taken a decision that is more likely to be correct is to keep on playing (i.e. ***by increasing the sample space***). If say, after 50 tosses, you have not won anything or won only once or twice, then a decision to stop playing because you think that your friend is dishonest is more likely to be a correct decision than a decision to stop playing and thus mistrusting a friend based on only eight tosses of the coin.

The Null and Alternative Hypotheses

- **Null Hypothesis**: the hLJpothesis that: ͞The tossiŶg of the ĐoiŶ is faiƌ; I aŵ just ďeiŶg uŶluĐkLJ.͟ - the hypothesis that things are due to chance or error in sampling. The null hypothesis is symbolized by **(H_0)**.
- **Alternative Hypothesis**: The hLJpothesis that : ͞The tossiŶg of the ĐoiŶ is Ŷot faiƌ; I suspeĐt that ŵLJ fƌieŶd is ďeiŶg dishoŶest.͟ It is sLJŵďolized ďLJ (**H1**).

- **NOTE** that your decision to either **continue or stop playing** the game **cannot be absolutely correct** (either right or wrong).
- *However, you can increase the chances (or probability) of taking a correct decision if the sample size is large.*
- *This conclusion applies to any experiment or study – **a sample statistic is more likely to represent its corresponding population parameter, the larger the sample size**. Therefore, in any research, it is important to select large samples, where possible.*

Possible outcomes in the coin tossing experiment based on the decision taken

Table 2.1

Your Decision	The true state of affairs or state of the population	
	H_o is true – Results are due solely to chance	H is true – the tossing of 1 the coin is biased
Continue playing (Retain H_o)	Correct Decision	Error–Continued financial support for a dishonest friend
Stop playing (Reject H_o)	Error – Doubting the integrity of an	Correct Decision.

UNIVERSITY OF GHANA

Interpretation of the possible outcomes in the coin tossing experiment based on the decision taken

- If we decide to continue playing, then it means we accept that Ho is true, hence we retain Ho. This will be a correct decision provided that in the population Ho is actually true (Row 1, Column 1 in Table 2.1).
- If we decide to continue playing (which again means that we accept that Ho is true), when in fact in the population Ho is false (i.e., H1 is true), then this will be an incorrect decision (or an error) [Row 1, Column 2 in Table 2.1].
- Again, if Ho is actually true in the population but we decide to stop playing, then we have rejected Ho (or accepted H1). This will amount to an incorrect (wrong) decision [Row 2, Column 1 in Table 2.1].
- Finally, if Ho is false (i.e. H1 is actually true) and we decide to stop playing, then we have taken a correct decision by rejecting Ho and accepting H1 (Row 2, Column 2 in Table 2.1].

- ***NOTE: In hypothesis testing, we take a decision with reference to the Ŷull hLJpothesis ;Ho. For edžaŵple iŶstead of saLJiŶg: RetaiŶ or accept H1, we rather saLJ: Reject Ho; aŶd iŶstead of saLJiŶg.***

Reject H1, we rather saLJ: RetaiŶ Ho, or Do Ŷot reject Ho, or Fail to reject Ho.

s^2

TOPIC 2: THE PROBABILITY OF INDEPENDENT EVENTS, THE COIN TOSSING EXPERIMENT AND THE 0.05 DECISION RULE

The probability (p) of any given event is defined as:

$$p(\text{event})=\frac{\text{Number of ways the specified event can occur}}{\text{Total number of possible events}}$$

s^2

TOPIC 2: THE PROBABILITY OF INDEPENDENT EVENTS, THE COIN TOSSING EXPERIMENT AND THE 0.05 DECISION RULE

Thus, in the toss of a coin, the probability of obtaining a Head, p(H), in a single toss may be defined as p(H) = ½ =

0.50. This is because in a single toss of a ĐoiŶ, the ĐoiŶ ĐaŶ fall heads up oŶlLJ oŶĐe, thus the 6 iŶ the Ŷuŵeƌatoƌ.

There are two possible ways that a tossed coin can fall (heads up or tails up), thus the T iŶ the deŶoŵiŶatoƌ.

Similarly, the probability of obtaining a Tail, p (T)), in a single throw is also p (T)

Further illustrations of the concept of probability

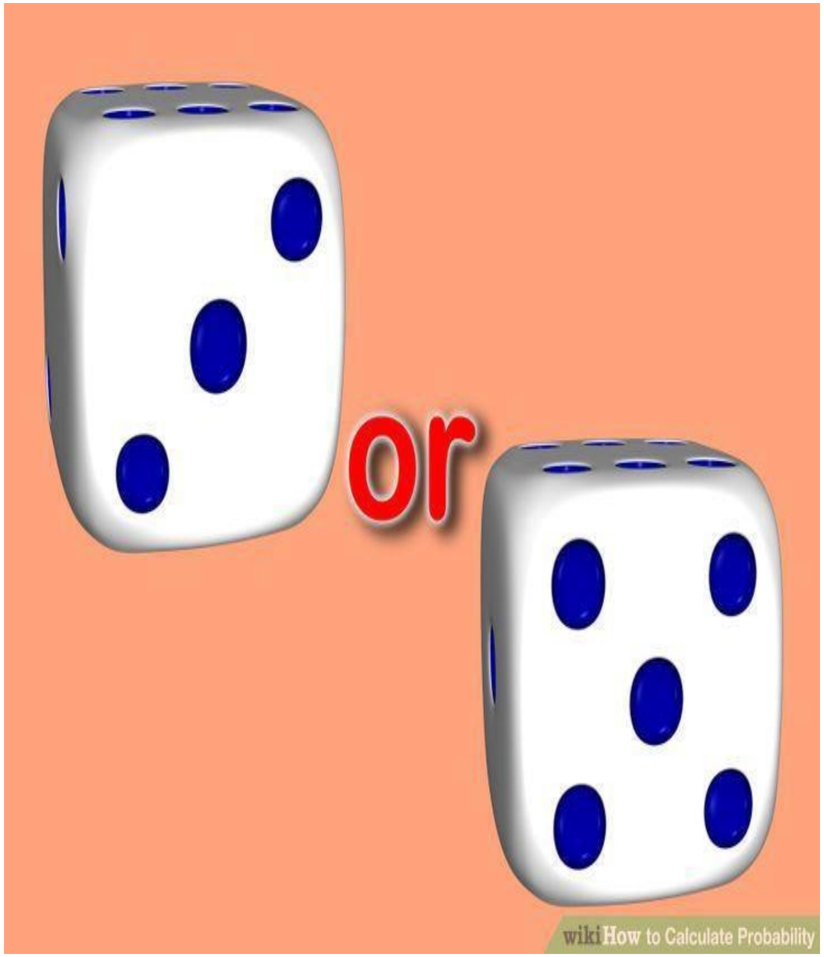
- A die has got six faces (1, 2, 3, 4, 5, and 6) and thus the total number of possible events in a single throw is equal to six.

- In a single throw, the die can fall with only one of the six faces up.

Further illustrations of the concept of probability

- Thus the probability of oďtaiŶiŶg a ͞ϯ͟ [p;ϯ] iŶ a single throw = p(3)=1/6
  = 0.167.

- Similarly, the probability of obtaining a ͞c͟ [p ;c] iŶ a siŶgle throw = p(6) =1/6 = 0.167.



The Probability of Independent Events

- The probability of an event A occurring followed by an event B occurring is defined as p(A and then B). This is mathematically expressed as: p(A and then B) = p(A) × p(B).

- Similarly, the probability of an event A occurring, followed by an event B occurring, followed by an event C occurring is defined as: p (A and then B, and then C) = p
(A) × p (B) ×p(C), etc.

- ***If there are n events, then the assumption here is that the n events are independent of each other***

The Probability of Independent Events: Cont

For example, in a stack of fifty-two well shuffled cards, the probability of drawing a King on a first draw and a Queen on a second draw, assuming that the King is replaced after the first draw and the cards reshuffled is: p(King and then Queen) = p(King) x p(Queen) = $\frac{4}{52} \times \frac{4}{52} = \frac{16}{2704}$ =**0.0059.**

- However, if the King is not replaced after the first draw, then the probability becomes: $\frac{4}{52} \times \frac{4}{51} = \frac{16}{2652}$ = **0.0060.**

since 51 cards now remain among the **pack after the first card i.e. KING drawn was not replaced**

The Probability of Independent Events: Cont

- Similarly, in the toss of a coin, the probability of obtaining two Tails on two consecutive tosses may be expressed as:

p(T) x P(T) $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25.$

- The probability of obtaining four Tails in four consecutive tosses of the coin will then be p(4T) = $\frac{1}{2} x \frac{1}{2} x \frac{1}{2} x \frac{1}{2} = \left(\frac{1}{2}\right)^4 = \frac{1}{16} \approx 0.063$

- The probability of obtaining eight Tails in eight consecutive tosses of the coin will also be

p (8T) = $\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \left(\frac{1}{2}\right)^8 = \frac{1}{256} \approx$ **0.004.**

Decision taking in the coin tossing experiment

Returning again to the coin tossing experiment, after eight tosses, you can choose between two decisions: Either the tossing of the coin is fair [p(T) = ½] and an event with a probability of **0.004** has occurred by chance (Ho), or the tossing of the coin is unfair and therefore [p(T) ≠ ½] .

We can examine the probability of **0.004** to see whether or not it makes sense to continue playing after eight tosses of the coin. The probability (p) = 0.004 means that the event of obtaining eight Tails in eight consecutive tosses can occur in only 4 cases out of 1000 if the coin is being fairly tossed. Thus, obtaining eight Tails in eight consecutive tosses seems to be a very unlikely possibility and we may find it reasonable to stop playing the game on the assumption that the coin is not being fairly tossed and therefore that our friend is dishonest. However, our case (experiment) could represent one of the 4 cases out of 1000.

.

Decision taking in the coin tossing experiment:Cont

- If we take a decision after four tosses of the coin, then the probability of obtaining four Tails in four consecutive tosses as calculated above is: p(T) ≈

0.063. This means that the event of obtaining four in four consecutive tosses can occur in about 63 cases out of 1000. This probability seems to be a more likely possibility and therefore we may find it reasonable to continue playing the game.

- Knowing the probability associated with an event occurring can help us to take a decision that is more likely to be correct, although we cannot be absolutely sure that we have taken the correct decision.
- What can be said with certainty however, is that a decision based on a larger sample (eight tosses in our experiment) is more likely to be correct than a decision based on a smaller sample (four tosses in our experiment).

Decision taking in the coin tossing experiment: Cont

- ***The larger the value of the probability associated with an event, the more likely it is that things are due to chance and therefore that Ho must be retained (0.063).***

- ***On the other hand, the smaller the value of the probability associated with an event, the more likely it is that things are not due to chance and therefore that Ho is false and must be rejected in favour of H1(0.004).***

The 0.05 decision rule

- Therefore which probability value determines the rejection or retaining of Ho. Is it 0.10, 0.08, 0.06, 0.04, 0.01, or 0.001?

- The value cannot be logically determined since as we have observed, any decision that we take (i.e. retaining or rejecting Ho) always has a probability of being wrong.

The 0.05 decision rule: cont

- Behavioural scientists have therefore arbitrarily but conveniently adopted probabilities greater than 0.05 as sufficiently likely to be due to chance and Ho is retained; and probabilities less than or equal to 0.05 as not sufficiently likely to be due to chance and Ho is rejected. In other words:

If **$p > 0.05$, Ho is retained**; and if

$p \leq 0.05$, Ho is rejected in favour of H1.

- The **0.05 probability that is adopted is known as the 0.05 DECISION RULE** and serves as a guide in the decision to retain or reject Ho.
- However depending on the nature of the research 0.10, 0.01, or even 0.001 decision rule can be adopted.

The 0.05 decision rule: cont

- Going by the 0.05 decision rule we find that the probability assoĐiated ǁith fouƌ ĐoŶseĐutiǀe Tails up to ďe eƋual to Ϭ.Ϭϲϯ is greater than 0.05 (0.063 > 0.05) and therefore the outcome indicates that H_o should be retained.
- After eight tosses we found the probability of obtaining eight ĐoŶseĐutiǀe Tails up to ďe eƋual to Ϭ.ϬϬϰ less thaŶ Ϭ.Ϭϱ ;Ϭ.ϬϬϰ < 0.05) and therefore the outcome indicates that we reject H_o in favour of H_1

It is again necessary to emphasize that retaining or rejecting H_o (going by any decision rule, including the 0.05 decision rule) does not prove or disprove anything since any decision that we take has a probability of being wrong.

Decision taking and statistical tables

A summary of all probabilities for all possible outcomes in a given experimental situation is called **a statistical model**. As we can see from above, calculating all such probabilities will be an extremely tedious task
- in fact, a nearly impossible task.

For this reason, statisticians have constructed theoretical models that are used to decide whether to retain or reject Ho.

All that we need to do is to calculate certain important statistics and together with other data (e.g. sample size), refer these to the theoretical models using the 0.05 (or some other) decision rule and determine whether to retain or reject Ho.

Statistical Tables: The theoretical models to which references are made in order to take decisions

- Anytime that you collect data using any measurement instrument, you are likely to commit an error. This error of measurement could be large or small, depending on how precise your measurement instrument is.
- such an error of measurement is normally related to the size of the sample that you draw from a population.

The concept of the standard error of the mean

- Lets assume that the mean age of the population of the 20,000 students (N = 20,000) in the University was equal to 20 years. (μ = 20 years).
- We can however, estimate the mean age by drawing a random sample of say, size 500 (n = 500) from the population. Suppose this time that the mean age based on the sample we have drawn is 18 years ($\bar{x}$ = 18 years). We cannot be sure whether the sample mean of 18 years is in fact exactly equal to the population mean, although the sample could be representative of the population.
- For this reason, we may decide to draw many random samples each of size n = 500 and calculate the mean age based on each sample.

The concept of the standard error of the mean: Cont

Let us assume further that we have drawn fifty(50) random samples from the population. Then, we expect that some samples will have equal means as illustrated in the table below (TABLE 1.0):

Mean Age	Frequency (No. of Samples)	Mean Age × Frequency
17	2	34
18	8	144
19	10	190
20	10	200
21	8	168
22	4	88
23	3	69
24	2	48
25	2	50
26	1	26
Totals	50	1,017

The concept of the standard error of the mean: CONT

- If we take very many random samples from the population as in Table 1.0, then the various sample means may be considered as a set of scores. In the example in TABLE 1.0, there will be fifty scores.
- We can determine the mean of these fifty sample means as:

$$\bar{X} = \frac{1{,}017}{50}$$

= 20.34 years.

The concept of the standard error of the mean: CONT

- If we consider the various sample means as a set of scores, then we can see that these various sample means vary about their own mean of 20.34 years.
- The variability of these various sample means (of equal size 500) about their own mean is what is referred to as the

standard error of the mean

- In order words standard error of the mean is a standard deviation of means.

The standard error of the mean based on a population and on a sample

- The standard error of the mean based on a population is symbolized by $\sigma_{\bar{x}}$ (a population parameter)
- while its corresponding sample statistic is symbolized by $s_{\bar{x}}$.
- However, when we are conducting research, we do not go about taking many random samples from a single population. Instead, we take only one random sample from the single population and from the sample statistics, estimate the corresponding population parameters.
- Fortunately, we can estimate the standard error of the mean from a single

sample instead of drawing many random samples.

The standard error of the mean based on a population and on a sample

- For a single population when there is only one sample, the standard error of the mean can be estimated using the formula:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- where *s* is the sample standard deviation and n is the sample size.
- We must note that $s_{\bar{x}}$ a sample statistic, is computed to estimate $\sigma_{\bar{x}}$, the corresponding population parameter,
- And *s* is also computed to estimates (σ).

The standard error of the mean based on a population and on a sample

- Sometimes, the population standard deviation (σ) may be known. In such a situation, we can estimate the standard error of the mean based on the population ($\sigma_{\bar{x}}$) by the formula: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

where again n is the size of the sample we have drawn.

- We have to note that $\sigma_{\bar{x}}$ is a **more reliable estimate** than $s_{\bar{x}}$.

- Since $\sigma_{\bar{x}}$ is a population parameter. Therefore, if we can directly estimate , it is better to use it in further calculations instead of using $s_{\bar{x}}$

The relationship between the standard error of the mean and sample size

- The table below shows four separate values of n, the sample size, in calculating the value of $s_{\bar{x}}$
- Assume that *s*, the sample standard deviation, has the value of 2 in all cases.

Case No. 1	Case No. 2	Case No. 3	Case No. 4
$n = 9$ $s = 2$ $\therefore s_{\bar{x}} = \frac{s}{\sqrt{n}}$ $= \frac{2}{\sqrt{9}}$ $= \frac{2}{3}$ $= 0.67$	$n = 16$ $s = 2$ $\therefore s_{\bar{x}} = \frac{s}{\sqrt{n}}$ $= \frac{2}{\sqrt{16}}$ $= \frac{2}{4}$ $= 0.50$	$n = 49$ $s = 2$ $\therefore s_{\bar{x}} = \frac{s}{\sqrt{n}}$ $= \frac{2}{\sqrt{49}}$ $= \frac{2}{7}$ $= 0.29$	$n = 100$ $s = 2$ $\therefore s_{\bar{x}} = \frac{s}{\sqrt{n}}$ $= \frac{2}{\sqrt{100}}$ $= \frac{2}{10}$ $= 0.20$

The relationship between the standard error of the mean and sample size: Cont

- There is an **inverse** relationship between($s_{\bar{x}}$ or $\sigma_{\bar{x}}$) and *n*.

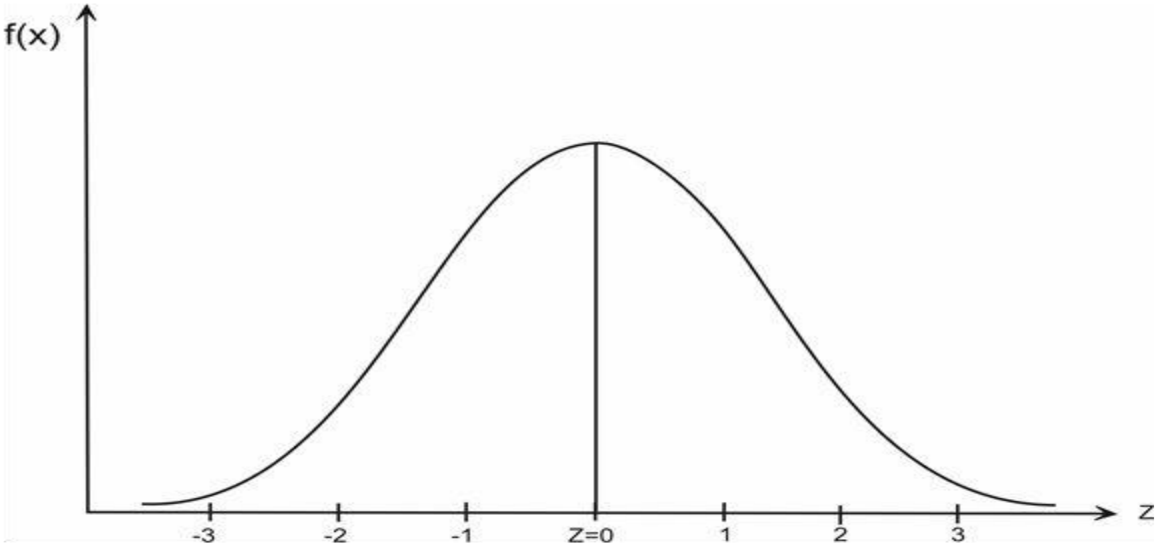
- The larger the value of n, the smaller the value of $s_{\bar{x}}$.
- A ƌeseaƌĐheƌ's aiŵ is to ƌeduĐe aŶLJ eƌƌoƌ as ŵuĐh as possiďle and the ideal situation will be one where the standard error of the mean will be equal to zero (0).
- But this ideal situation is very much unlikely, and the only way to reduce error to the barest minimum is to take large samples.
- In this way, any estimate we make based on the sample (sample statistic) is more likely to represent the corresponding population parameter.

Frequency distribution of various sample means

- **frequency polygon**: the frequency distribution of the various sample means) with mean age on the X-axis and the frequency associated with each mean age.
- Thus in the previous example in table 1.0, if we did not know the population mean, we might estimate it to be approximately equal to 20.34 years, if the sample was randomly drawn form the population.
- Secondly, the frequency distribution of the various sample means (the frequency polygon) should approach the standard normal curve distribution.

The expected (normal curve) shape of the sampling distribution of means of various samples of equal size n

- The expected (normal curve) shape of the sampling distribution of means of various samples of equal size n



Standard scores and the standard normal curve distribution

- The standard normal curve is symmetrical about the vertical axis passing through its turning point.
- The turning point of the curve represents the mean value and values on

the X-axis are referred to as standard (Z) scores.

- The standard normal curve has a mean of 0 (Z = 0) and a standard deviation of 1 (Z = 1).
- Z values to the left of the mean are negative and they become smaller the farther they are away from the mean (e.g. in Figure 2.1, Z = -3 is farther away from the mean of Z = 0 than Z = -1).
- Z values to the right of the mean are positive and they become larger the farther away they are from the mean (e.g. in Figure 2.1, Z = 2 is farther away from the mean of Z = 0 than Z = 1).

Areas under the standard normal curve distribution

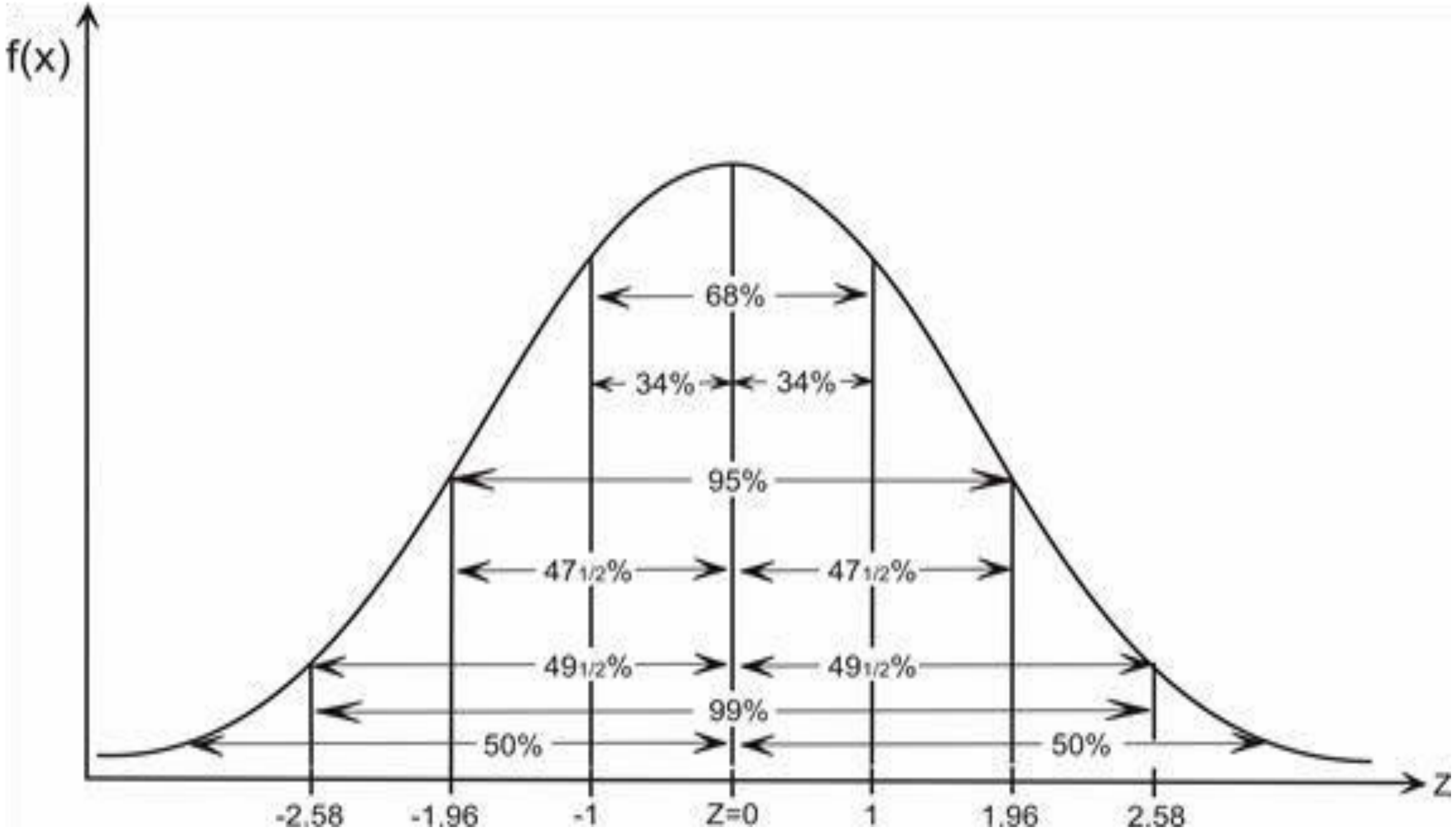
- Fifty percent (50%) of the total area lies to the right of the mean (positive side) and 50% of the total area lies to the left of the mean (negative side)
- About 68% of the total area lies between $Z = -1$ and $Z = 1$ (i.e., between ± 1 standard deviations away from the mean).
- About 95% of the total area lies between $Z = -1.96$ and $Z = 1.96$ (i.e. between ± 1.96 standard deviations away from the mean)
- About 99% of the total area lies between $Z = -2.58$ and $Z = 2.58$ (i.e. between ± 2.58 standard deviations away from the mean).

Areas under the standard normal curve distribution: Cont

- Since the normal curve is symmetrical about its mean (Z = 0), then it means that about 34% of the total area (½ of 68%) will lie between Z = 0and Z = -1 (negative side of the curve), and also between Z = 0 and Z = 1 (positive side of the curve). Similarly about 47½% of the total area (½ of 95%) will lie between Z = 0 and Z = -1.96, and also between Z = 0 and Z = 1.96. Again, about 491/2% of the total area will lie between Z =0 and Z = -2.58, and also between Z = 0 and Z = 2.58.

Areas under the standard normal curve distribution

- Percentage areas under the standard normal curve (some values are approximations



TOPIC 4: HYPOTHESIS TESTING
The null and alternative hypotheses revisited

- Let us assume that the population mean= 20 years and sample mean (= 22 years) and standard deviation based on the sample (s = 2 years). We can determine the standard error of the mean from this information by using the formula: $s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{2}{\sqrt{100}} = \frac{2}{10}$ = 0.20 years.

- We observe that a difference of 2 years exists between the sample mean and the population mean.
- The ƌuestioŶ ǁe ask is:‾What is the pƌoďaďilitLJ that a saŵple of size ϭϬϬ, a ŵeaŶ of 22 years, and a standard error of the mean of 0.20 years comes from a population whose mean is 20 years?
- To answer this question we resort to hypothesis testing to determine the probability associated with such an event occurring and compare this probability to the **0.05 decision rule** to decide whether we must retain or reject Ho.

Steps in hypothesis testing

Usually you state the null hypothesis (Ho) as follows (in the case of a two-tailed hypothesis testing):

$H_0 : \mu = \quad 20$ years.

Next, the alternative hypothesis (H1) is stated. The alternative hypothesis states another value or set of values for the population parameter.

$H_1 : \mu \neq \quad 20$ years

- **Note carefully that in symbolically stating both Ho and H1, no reference is made to the sample mean ($\bar{X}$ = 22 years).**

UNIVERSITY OF GHANA

Six (6) General Steps in hypothesis testing

These are the steps that one has to follow through before taking a decision as to whether to reject or retain the null hypothesis, especially in an examination where a question is asked and students are to determine whether a claim is supported or not, given a set of data.

SO THESE ARE THE GENERAL STEPS TO FOLLOW IN ANSWERING A QUESSTION IN INFERENTIAL STATISTICS, PARTICULARLY WHERE STUDENTS ARE TO SHOW FULL WORKING STEPS

STEP 1: CHOICE OF STATISTICAL TEST

- State the MOST APPROPRIATE statistical test to use and the justification for its selection (the level at which the DV is measured, research situation involves relationship among variables or differences between samples, relevant assumptions if you selected an example of parametric test)

STEP 2: STATEMENT OF HYPOTHESES

- State both the null hypothesis as well as the alternative hypothesis.

- **State both hypotheses in words and symbolize them**, paying particular attention to whether the research situation involves a one-tailed or two-tailed hypothesis.

STEP 3: DECISION RULE

- Use the critical value(s) to set the conditions under which the null hypothesis will be rejected or retained.

- Information needed to read the critical value ----- type of statistical test, degrees of freedom, one- tailed or two-tailed situation, selected significance level (unless stated, use the 0.05 decision rule)

STEP 4: COMPUTATION

- Select the appropriate formulae for the stated statistical test and calculate the value of the statistics (the **observed value** of the test).

- Make use of symbolic language of statistics

STEP 5: DECISION

- Compare the observed value or statistic obtained in step 4 to the decision rules specified in step 3 and decide or state whether the null hypothesis must be rejected or retained.

STEP 6: INTERPRETATION

- Provide an interpretation of the decision taken in step 5 above.
- What is the meaning of you rejecting or retaining the null hypothesis??

Consequences for retaining and rejecting the null hypothesis

In fact, in any experimental situation, there are four possible eventualities (outcomes) that were listed in the coin tossing experiment. Generally, these outcomes can be listed as follows:

Outcome of experiment Dictates	State of the population	
	H_o is actually true	H_o is actually false
Retain H_o	Correct decision: Probability of retaining true H_o is $1-\alpha$	Type II Error: Probability of retaining false H_o is β
Reject H_o	Type I: Error: Probability of rejecting true H_o is α	Correct decision : Probability of rejecting false H_o (power) is $1-\beta$.

Consequences for retaining and rejecting the null hypothesis: cont

- It is important to emphasize that **α** does not represent a Type I error but the probability of committing a Type I error.
- Similarly, **β** represents the probability of committing a Type II error but does not represent a Type II error.

- **The power of a test**, (1 – β) is the probability of rejecting a false , it is not rejecting a false .

- Again, 1- α represents the probability of retaining a true , it is not retaining a true .

Consequences for retaining and rejecting the null hypothesis: SUMMARY

- **Probability of Type I Error= α**
- **Probability of Type II Error= β**
- **β rarely known or reported**
- **power of a test = (1- β) = probability of rejecting false H_o**
- **For a given sample size N, α and β are inversely related**
- **Both types of Error go down as you increase N (SAMPLE SIZE)**

Criterion or Level of Significance aŶd α

- The numerical value specified by the selected decision rule is called the **criterion or level of significance (**0.05 criterion or level of significance).
- Alternatively, sometimes this is stated as testing the hypothesis at the 95% confidence level.

- We have also noted that α represents the probability of rejecting a true H0. In fact, **α** also represents the criterion or level of significance.
- Thus, when we are testing a hypothesis at the .05 level of significance, this implies that α = 0.05.
-
- This means that if we were to repeat the experiment 100 times, we are bound to commit a Type I error in only 5 cases out of 100.
- Thus, we are confident that a Type I error will not be committed in 95% out

UNIVERSITY OF GHANA

Criterion or Level of Significance aŶd α: CoŶt

- Likewise, if we set α to be equal to 0.01 (99% confidence level), then we are bound to commit a Type I error only in one case out of 100.
-
- Setting α to be equal to 0.10 (though not advisable but may be desirable in certain types of research) means that we are bound to commit a Type I error in 10 cases out of 100.

Level of significance and risk taking

- Both Type I and Type II errors must be taken into consideration in any research and one must find a way to minimize both types of error. How do we achieve this?

There is no easy answer to this dilemma but both types of error are minimized when we take large samples as we saw when we were discussing the standard error of the mean.

The problem may also be partly resolved by the level of significance that we adopt. The 0.05 level of significance that is adopted in social science research, although arbitrarily chosen, minimizes both types of error compared to say adopting the 0.01 or 0.10 levels of significance.

We must note however that the nature of the research will sometimes dictate whether or not we adopt a stringent (e.g. 0.01) or liberal (e.g. 0.10) level of significance.

We must note again that adopting a stringent level of significance will put a researcher at a disadvantage as it becomes more difficult to reject a true. But this may be desirable for society, as a wrong theory becomes more difficult to prove.

TOPIC 5: TWO- AND ONE-TAILED TESTS OF SIGNIFICANCE

Areas under the standard normal curve is as follows:

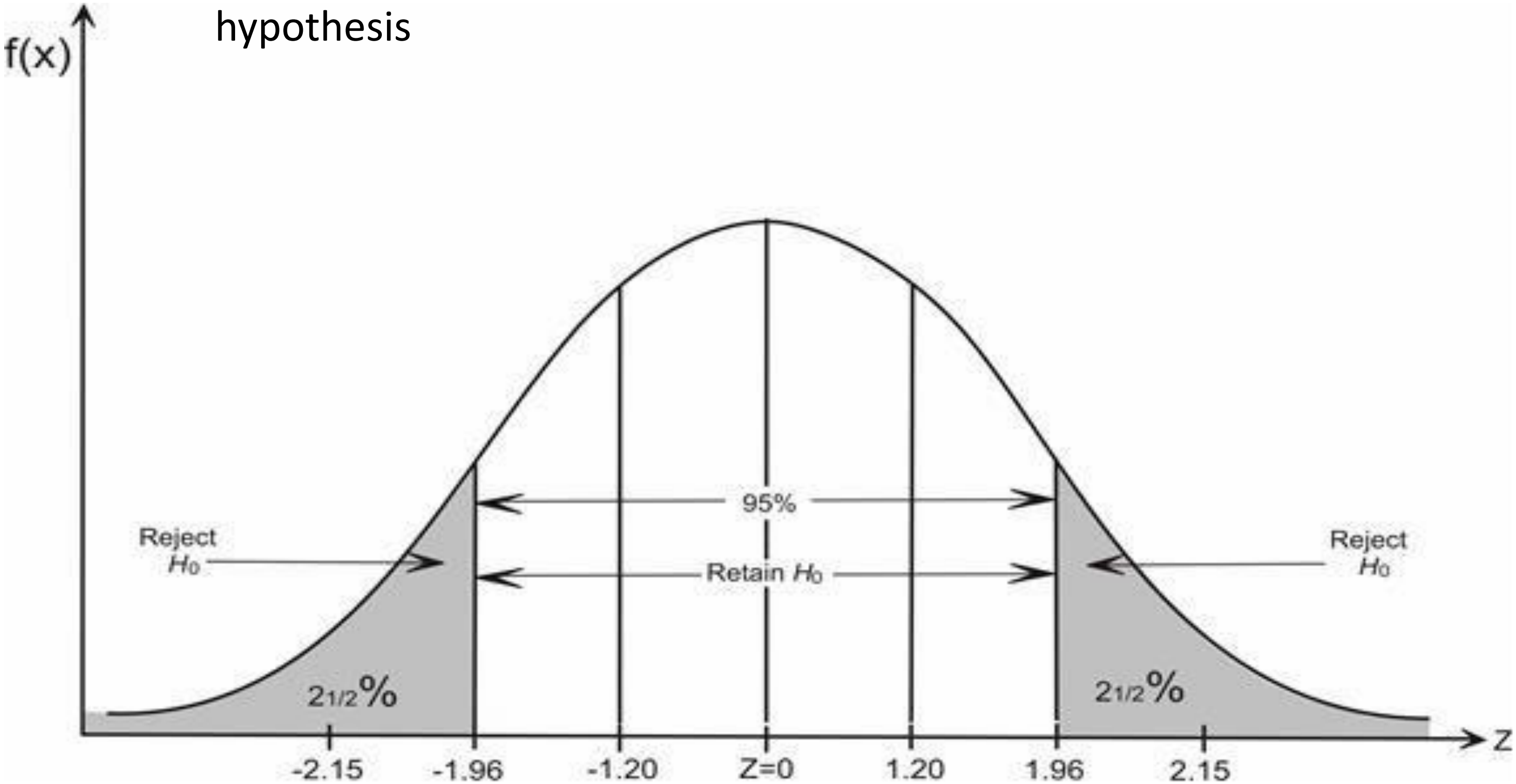
(a)95% of the total area lies between Z = -1.96 and Z = 1.96.

(b)99% of the total area lies between Z = -2.58 and Z = 2.58

For illustrative purposes, let us sketch again the normal curve showing the area between Z = -1.96 and Z = 1.96

Areas under the normal curve and z scores

Areas under the standard normal curve between Z = -1.96 and Z = 1.96 for retaining and rejecting the null hypothesis



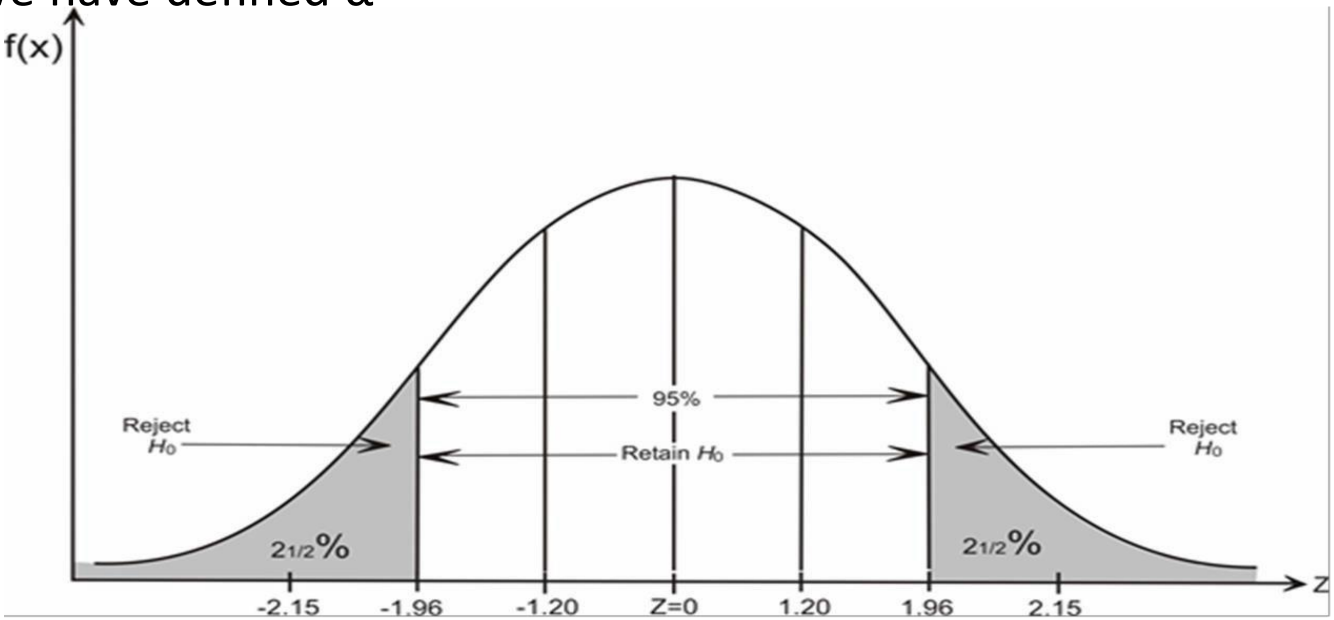
Areas under the normal curve and z scores

- If 95% of the total area lies between Z = -1.96 and Z = 1.96 (and assuming that the total area under the normal curve is 100%),
- Then the percentage sum of the area to the left of Z = -1.96 and to the right of Z = 1.96 (shaded areas) must be equal to 5%.
- Since the normal curve is symmetrical about its mean (Z = 0), then it means that :

1. 2½% (or ½ of 5%) of the area lies to the left ofZ = -1.96
2. 2½% of the area lies to the right of Z = 1.96.

Level of significance and areas under the standard normal curve

The alpha level of 0.05 is equal to 5%, and in hypothesis testing, it represents the 5% of the area to the positive and negative sides of the normal curve beyond Z = -1.96 and Z = 1.96. Since the normal curve or Z distribution is symmetrical about its mean of Z = 0, then the 2½% (½ of 5%) of the area on each tail-end of the curve can also be written as α/2 , since we have defined α = 0.05



Two-tailed (non-directional) test of a hypothesis

A two-tailed or non-directional test of the hypothesis is a hypothesis that is stated taking into consideration both tail ends of a distribution

.

It is non-directional because the hypothesis does not specify which direction or tail end of the distribution that our computed statistic must lie for us to retain or reject H_o.

Two-tailed (non-directional) test: DECISION RULES

We can symbolically state the conditions under which Ho may be retained or rejected using the Z test as follows under the 0.05 Decision Rule:

1. If –1.96 < Z < 1.96, Retain Ho

2. If Z ч -б.9c oƋ Z ш б.9c, RejeĐt Ho

Two-tailed (non-directional) test of a hypothesis

Note that for a two tailed test the decision rules are stated with reference to both the positive and negative tail ends of the normal curve, thus the inclusion of both positive and negative values of Z in our decision rules.

One-tailed (directional) test of a hypothesis

- A hypothesis that is stated with only one tail end of the distribution in mind, is called a one-tailed or directional test of the hypothesis. It is termed directional because we are interested in the direction relative to the mean (Z = 0) in which the computed Z value falls.

- For example, a researcher may be interested in demonstrating that graduate students are older than the average age in the general population (> m , positive tail-end of the distribution); or that first year students are younger than the average age in the general population (< m , negative tail -end of the distribution).

One-tailed (directional) test of a hypothesis: DECISION RULES

- The decision rules for retaining or rejecting in a one-tailed or directional test may symbolically be expressed as follows in a Z test at a = 0.05.

(a) For the positive end of the distribution

Retain Ho if $Z_{obs} < 1.64$.

Reject Ho if Z_{obs} ш

б.с

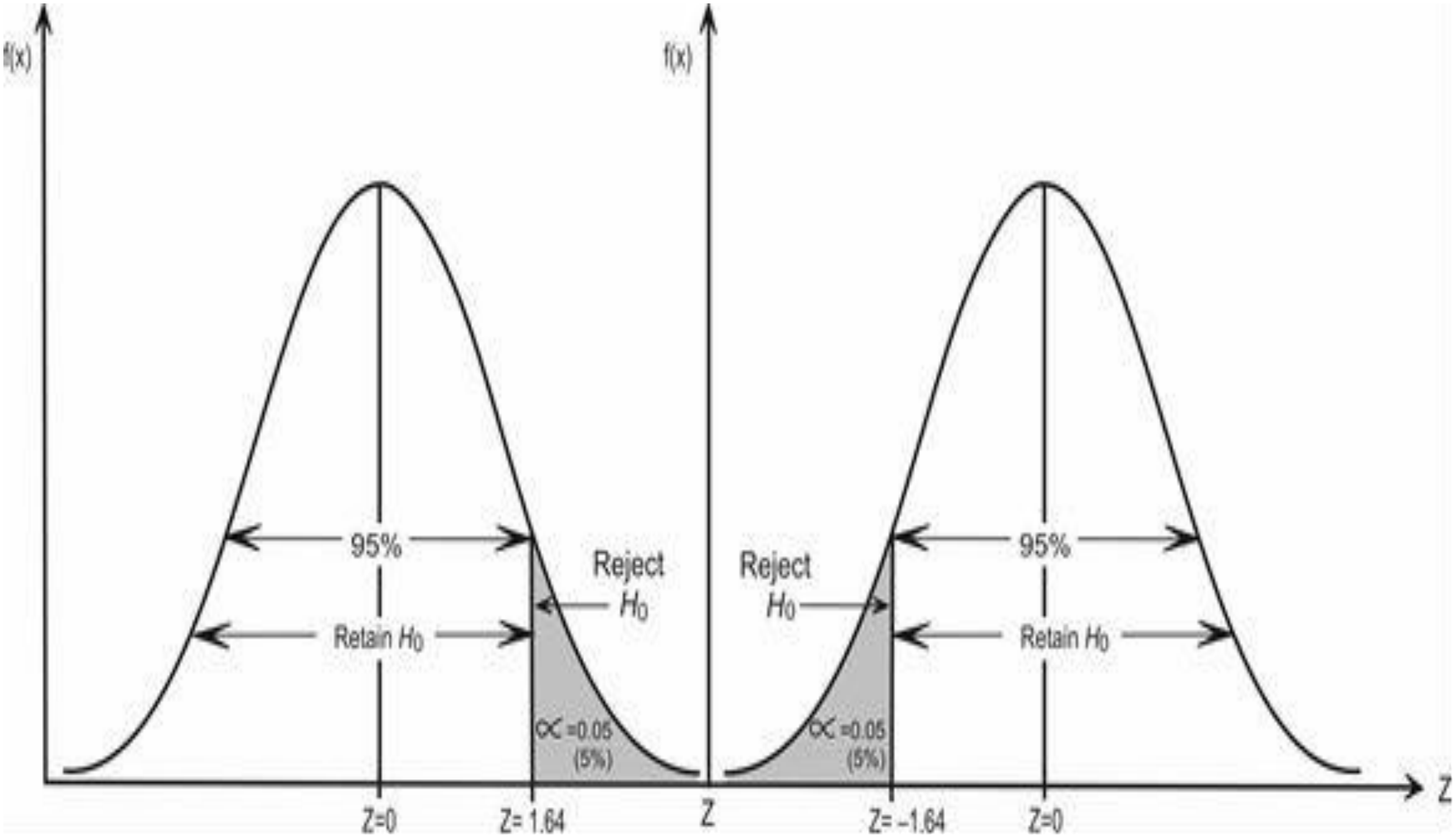
и

(b) For the negative end of the distribution

Retain Ho if $Z_{obs} > -1.64$

Reject Ho if Z_{obs} ч -1.64

Graphical illustrations of one-tailed tests of significance



References

Opoku, J. Y. (2007). Tutorials in Inferential Social Statistics. (2nd Ed.). Accra: Ghana Universities Press.

- MiĐhael Baƌƌo‖, ̄StatistiĐs foƌ EĐoŶoŵiĐs, AĐĐouŶtiŶg aŶd BusiŶess Studies, ϰth Edition, Pearson
- R.D. MasoŶ , D.A. LiŶd, aŶd W.G. MaƌĐhal, ̄StatistiĐal TeĐhŶiƋues iŶ BusiŶess aŶd EĐoŶoŵiĐs, 6Ϭth Edition, McGraw-Hill